cortex 141 (2021) 311–321

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex


Research Report

# The time course of speaker-specific language processing

Leon O.H. Kroczek [a,*] and Thomas C. Gunter [b]

[a] Department of Clinical Psychology and Psychotherapy, University of Regensburg, Germany
[b] Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

ARTICLE INFO

*Article history:*
Received 17 September 2020
Reviewed 5 December 2020
Revised 27 April 2021
Accepted 27 April 2021
Action editor Sonja Kotz
Published online 17 May 2021

*Keywords:*
Speaker identity
Language use
Syntax
Prediction
ERP

ABSTRACT

Listeners are sensitive to a speaker's individual language use and generate expectations for particular speakers. It is unclear, however, how such expectations affect online language processing. In the present EEG study, we presented thirty-two participants with auditory sentence stimuli of two speakers. Speakers differed in their use of two particular syntactic structures, easy subject-initial SOV structures and more difficult object-initial OSV structures. One speaker, the SOV-Speaker, had a high proportion of SOV sentences (75%) and a low proportion of OSV sentences (25%), and vice-versa for the OSV-Speaker. Participants were exposed to the speakers' individual language use in a training session followed by a test session on the consecutive day. ERP-results show that early stages of sentence processing are driven by syntactic processing only and are unaffected by speaker-specific expectations. In a late stage, however, an interaction between speaker and syntax information was observed. For the SOV-Speaker condition, the classical P600-effect reflected the effort of processing difficult and unexpected sentence structures. For the OSV-Speaker condition, both structures elicited different responses on frontal electrodes, possibly indexing effort to switch from a local speaker model to a global model of language use. Overall, the study identifies distinct neural mechanisms related to speaker-specific expectations.



## 1. Introduction

Communication is not just words. In a typical communicative situation, a vast number of information is available to the listener which has to be processed in order to arrive at a complete representation of what is being communicated. Besides the semantic and syntactic information in the language input itself, additional information can be provided by gesture (Holle & Gunter, 2007) or prosody (Hellbernd & Sammler, 2018), but also by contextual factors, like setting (Hay & Drager, 2010) or speaker identity (Brown-Schmidt, Yoon, & Ryskin, 2015; Lattner & Friederici, 2003; Van Berkum, 2008). With all this information available to the

* *Corresponding author*. Department of Clinical Psychology and Psychotherapy, University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany.
E-mail address: leon.kroczek@ur.de (L.O.H. Kroczek).
https://doi.org/10.1016/j.cortex.2021.04.017

listener, the question arises how such information can be used to constrain language processing.

Characteristic language use of a particular speaker allows listeners to use regularities of the language input to generate internal models related to a speaker's language use and to further use these models to build expectations regarding upcoming language (for a general mechanism see Kuperberg & Jaeger, 2015). Incorporating such expectations can increase the computational efficiency in the language system (Fine, Jaeger, Farmer, & Qian, 2013; Hale, 2001; Levy, 2008). Previous studies have found evidence that listeners use (speaker-specific) expectations in language comprehension (Brothers, Dave, Hoversten, Traxler, & Swaab, 2019; Federmeier, 2007; Fine et al., 2013; Hanulíková, van Alphen, van Goch, & Weber, 2012; Kroczek & Gunter, 2017). The exact timing of expectation processing in language comprehension remains under debate. For instance, previous studies demonstrated early effects of syntactic computation in language comprehension (Friederici, 2002; Gunter, Friederici, & Schriefers, 2000) while the influence of expectations has been shown in a late stage of processing (Hanulíková et al., 2012). Such findings raise the question whether expectation processing is fundamental in language comprehension (Huettig, Mani, & Huettig, 2015; but see Kuperberg & Jaeger, 2015). Note that, expectation processing does not necessarily exclude more stimulus-driven syntactic processing and vice-versa. A recent fMRI study, for instance, found distinct brain networks related to syntactic processing and speaker-specific expectancy processing (Kroczek & Gunter, 2020). In order to detail the temporal relation between speaker information and syntactic processing it is important to include high-grained information, as provided in EEG measures.

Models of the neuro-cognition of language describe different stages of processing when a sentence is encountered (Bornkessel & Schlesewsky, 2006; Friederici, 2002, 2011). With regard to syntactic structure processing, these stages have been related to different ERP components and typically divide into early and late stages of syntactic computations. Early stages have been related to a left-anterior negative ERP component (i.e., the LAN) that is elicited between 300 and 500 msec after a morpho-syntactic violation has been presented (Osterhout & Mobley, 1995). In addition, a late stage of processing is indexed by the P600 component, a positivity over centro-parietal electrodes that peaks after 600 msec, that has been related to syntactic processing in terms of re-analysis and integration (see Friederici, 2018). The P600 has further been linked to a more general representation of what is being communicated (Brouwer & Hoeks, 2013) and has been shown to be sensitive to information provided by gestures (Holle et al., 2012) or pragmatics (Regel, Coulson, & Gunter, 2010). It has also been argued that the P600 reflects domain general processing of deviants (Sassenhagen & Bornkessel-Schlesewsky, 2015; Sassenhagen, Schlesewsky, & Bornkessel-Schlesewsky, 2014). In summary, these different stages of language processing provide a useful framework to test the role of speaker-specific expectations in language comprehension.

In the current study we have implemented a paradigm, which allows listeners to generate expectations for a particular syntactic structure on the basis of speaker identity. Two speakers (one female, one male) were presented to the participants. These two speakers varied in the probability by which they used particular syntactic structures. One speaker had a high probability to produce sentences with a difficult Object–Subject–Verb (OSV) structure, whereas the other speaker had a high probability to produce sentences with an easy Subject–Object–Verb (SOV) structure. In an initial training session, participants were exposed to this particular speaker–syntax-coupling that allowed them to generate syntactic predictions on the basis of speaker identity. In a subsequent EEG session on the next day, we presented participants with the same speakers and the same syntactic structures while measuring the EEG. ERPs were time locked to the onset of the determiner of the first noun, as this was the position where the sentence's syntactic structure became clear to the listener. Depending on the role of predictions during syntactic processing, we hypothesized to find differential effects with regard to the timing of a possible interaction of speaker and syntax information. If predictions affect core processes of syntactic structure building, we would expect to find an interaction in an early stage of processing. In contrast, if speaker-specific predictions are used in creating a coherent representation of what is being communicated we hypothesized to find a modulation of late syntactic effects (e.g., a P600 effect) by speaker identity.

## 2. Methods

### 2.1. Participants

Thirty-seven German native-speaking adults were paid for participation. Due to excessive artifacts (i.e., less than 50% of trials in a condition after artifact correction), five participants were excluded from further analysis. The remaining 32 participants ($M_{age}$ = 24.91, $SD_{age}$ = 2.26, range 20–29 years, 16 female) were right-handed, had a mean laterality quotient of 94.9 ($SD_{LQ}$ = 7.2, Oldfield, 1971) and reported neither a hearing problem nor a history of neurological impairment. Sample size was chosen on the basis of previous studies with similar paradigms (Hanulíková et al., 2012; Kroczek & Gunter, 2020). All participants gave written consent following the guidelines of the Ethics committee of the University of Leipzig (Nr: 236-10-2382010), in accordance with the declaration of Helsinki.

### 2.2. Materials

A set of 360 German sentence-items was used in the study (240 items in the EEG experiment and 120 items in the training). All items consisted of combinations of a lead-In phrase (LP), two noun-phrases (NP), and a verb (V). Noun-phrases consisted of a determiner (DET) and a noun (N). The case information of the determiner indicated whether the noun-phrase was the subject or the object of the sentence. A nominative determiner ("der") marked a subject NP and an accusative determiner ("den") marked an object NP. Nouns were always male and singular and were selected to be plausible both as the subject and as object of the sentence. All verbs were transitive and did not have any semantic bias towards one of the nouns. Every item was used to create a Subject–Object–Verb (SOV) sentence structure and an Object–Subject–Verb (OSV) sentence structure. Additionally, two versions of every sentence were created, with subject-nouns and the object-nouns exchanged across versions. In summary, every item was used to generate four different sentences (Table 1).

The experimental sentences were spoken by two voice actors (one female, one male) and recorded with a sample rate of 44.1 kHz (Audacity v 2.0). A 50 msec period of silence was added at the beginning, at the end and at the onset of the first determiner of every sentence. Sentences were normalized using the Root Mean Square (RMS) of the amplitude. Auditory sentence stimuli had an average duration of 2681 msec (SD = 231 msec). The complete stimulus set consisted of 1920 audio files for the EEG experiment and 960 audio files for the training.

### 2.3. Experimental procedure

The experiment consisted of a training session and an EEG session on two consecutive days. For every participant a randomized list was created for both sessions and sentence items were randomly assigned into conditions of syntactic structure and speaker identity. Importantly, the training session on the first day was conducted to introduce the speaker–syntax coupling, where a particular syntactic structure was high-frequent for sentences produced by one speaker and low-frequent for sentences produced by the other speaker. The EEG session on the next day was then conducted to test the electrophysiological effects of this speaker–syntax coupling. This procedure, including separate training and test sessions on two consecutive days, has been shown to induce syntactic expectations on the basis of speaker identity in the listeners (Kroczek & Gunter, 2017).

Experimental procedure was identical between the training session and the EEG session. Participants listened to 480 sentence stimuli per session. There were 240 sentences per speaker. Importantly, the frequency of syntactic structure differed between speakers. One particular speaker, the so called SOV-*Speaker*, produced SOV structures in 75% of the sentences (180 trials) and OSV sentences in 25% of the sentences (60 trials). For the other speaker, the so called OSV-*Speaker*, this pattern was reversed, i.e., 75 % OSV sentences (180 trials) and 25% SOV sentences (60 trials). Sentence stimuli presented in the training session were not used in the EEG session. Crucially, the introduced speaker–syntax coupling always remained the same for the particular speakers between training session and EEG session. However, the assignment of the male and the female speaker as SOV and OSV speaker was counterbalanced across participants.

Stimulus presentation was controlled by Presentation ® (Neurobehavioral Systems) in both sessions. A single trial started with the presentation of a fixation-cross for 500 ms. Then the sentence stimulus was presented via (loud-) speakers (average duration 2681 msec). The fixation-cross remained on the screen during the presentation of the auditory stimuli and lasted 500 msec after the presentation of the sentence. Then a comprehension question was displayed on the screen. The question was related to the previous sentence stimulus and always had the form "*Was the* NP1/NP2 *greeted?*" (e.g., "*Was the man greeted?*"). Whether the question asked for NP1 or NP2 was randomized. Participants had 2000 msec to respond via a button-press with "*Yes*" or "*No*". Key assignment of the response options was balanced over participants. Immediately after the response or after 2000 msec a feedback stimulus was presented for 500 msec (correct response = "smiley", incorrect response = "frowny", miss = clock symbol). After the feedback, the next trial started. Every 120 trials there was a break of self-determined length.

The training session was conducted in front of a computer screen with no further restrictions. In the EEG session participants were placed in a dimly lit, electrical shielded room in front of a computer screen. Participants were asked to restrict movements and eye blinks during sentence presentation in order to avoid signal artifacts. The training session lasted about 50 min, while the EEG session had a duration of about 2.5 h including electrode application.

### 2.4. EEG recording and analysis

The EEG was recorded continuously using 64 Ag/AgCl electrodes mounted in an elastic cap according to the 10–20-system. Sternum served as ground. The EEG was amplified using a PORTI-32/MREFA amplifier (DC to 135 Hz) and digitized online at 500 Hz. Impedances were kept below 5 kΩ. EEG was referenced to the left mastoid during data acquisition. Bipolar EOG was measured horizontally and vertically.

Offline pre-processing was conducted using Matlab (v 8.6, MathWorks) toolbox Fieldtrip (Version: 20180501; Oostenveld, Fries, Maris, & Schoffelen, 2011). Artifact correction was performed using independent component analysis (Jung et al., 2000). For that reason, preprocessing was conducted in two steps. First, raw EEG was filtered using a windowed-sinc finite impulse response high-pass filter with a cut-off of 1 Hz (−6 dB, half amplitude, onepass-zerophase, Kaiser window, maximum passband deviation of .001; Widmann, Schröger, & Maess, 2015), re-referenced to the average and an independent component analysis was performed using runica with the infomax algorithm (Delorme & Makeig, 2004). In a next step, another dataset was created by filtering the raw EEG data using a windowed-sinc finite impulse response high-pass filter with

**Table 1 – The item (LP = "Heute hat"/"Today has", NP1 = "der Mann"/"the man", NP2 = "der Freund"/"the friend", V = "grüssen"/"to greet") was used to create sentences with two types of structures (SOV and OSV). Additionally two versions were created which differed with regard to their subject-nouns and object-nouns.**

| Structure | SOV | OSV |
|---|---|---|
| Version 1 | Heute hat der Mann den Freund gegrüsst.<br>*Today has the man$_{[Nom]}$ the friend$_{[Acc]}$ greeted.* | Heute hat den Freund der Mann gegrüsst.<br>*Today has the friend$_{[Acc]}$ the man$_{[Nom]}$ greeted.* |
| Version 2 | Heute hat der Freund den Mann gegrüsst.<br>*Today has the friend$_{[Nom]}$ the man$_{[Acc]}$ greeted.* | Heute hat den Mann der Freund gegrüsst.<br>*Today has the man$_{[Acc]}$ the friend$_{[Nom]}$ greeted.* |

a cut-off of .1 Hz (−6 dB, half amplitude, onepass-zerophase, Kaiser window, maximum passband deviation of .001) and re-referencing to the average of all electrodes. The ICA components from the first step were then projected on this new dataset. Components related to eye-blinks, eye-movements or muscle artifacts were removed from the EEG data. The ICA-corrected data were then again re-referenced to the linked mastoids and time-locked to the onset of the determiner of the first noun-phrase. Importantly, this is the position where the syntactic structure of a sentence is revealed. Epochs lasted from 200 msec prior to stimulus onset to 1000 msec afterwards. A 100 msec pre-stimulus baseline was applied. Finally, because trial numbers differed between conditions (e.g., 180 trials in the SOV-Speaker SOV condition and 60 trials in the SOV-Speaker OSV condition), only 60 out of 180 trials in the SOV-Speaker SOV and OSV-Speaker OSV conditions were entered into further analysis. This selection was conducted such that the analyzed trials from the frequent sentence conditions (i.e., SOV-Speaker SOV and OSV-Speaker OSV) had been presented directly before or after the corresponding trials form the infrequent conditions (i.e., SOV-Speaker OSV and OSV-Speaker SOV). This procedure was applied to reduce the impact of cognitive or attentional drifts across the course of the experimental session, while maintaining an equal signal-to-noise ratio between conditions.

Epochs with amplitudes exceeding ±100 μV were rejected from further analysis. The mean rejection rate was 10.83% (SD: 6.89). Only trials with a correct response in the comprehension task entered the analysis. On average there were 47.81 (SD: 5.31) trials in the SOV-Speaker SOV condition, 40.84 (SD: 10.49) trials in the SOV-Speaker OSV condition, 46.47 (SD: 6.3) trials in the OSV-Speaker SOV, and 41.84 (SD: 9.78) trials in the OSV-Speaker-OSV condition. Single subject averages were calculated for all four experimental conditions: SOV-Speaker SOV, SOV-Speaker OSV, OSV-Speaker SOV, OSV-Speaker OSV. For data visualization the grand average was calculated for all conditions by averaging across participants and a low-pass filter with a cut-off of 8 Hz was applied (see the supplementary material for a presentation of the ERP difference waves without a 8 Hz low-pass filter).

Statistical analysis was conducted by using cluster-based permutation tests implemented in the Fieldtrip software (Oostenveld et al., 2011). In order to investigate early and late effects, a time-window of 300–500 msec as well as 600–900 msec was selected based on existing literature (Friederici, 2002; Tanner & Van Hell, 2014). All channels were entered into the analysis. Correction for multiple comparisons was performed using cluster analysis based on Monte Carlo simulations (Maris & Oostenveld, 2007). In order to evaluate the interaction between *Speaker* and *Structure*, difference waves were first calculated by subtracting SOV sentence conditions from OSV sentence conditions for both speakers (SOV-Speaker, OSV-Speaker) separately. In a next step, the individual difference waves were entered into a cluster-based permutation analysis using t-tests for dependent samples ("statfun = depsamplesT"). Monte Carlo simulations were performed using N = 1000 randomizations. Clusters were defined based on cluster effect size ("clusterstatistic = maxsum", "minnbchan = 2", "correctm = cluster"). Main effects of *Speaker* and *Structure* were analyzed using the same procedure. In these cases, however, the single subject averages for each condition were entered into the analysis instead of difference waves.

### 2.5. *Behavioral analysis*

Statistical analyses of the behavioral data were conducted using the R environment (R Core Team, 2016) with packages *lme4* (Bates, Mächler, Bolker, & Walker, 2015), *car* (Fox & Weisberg, 2011), and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2016) installed.

Performance in the comprehension task was analyzed using a logit mixed-effects model for categorical response types (correct *vs* incorrect) and a linear mixed-effects model for the reaction times. Both models included fixed effects for the factors *Structure* (sum coded: SOV = 1, OSV = −1), *Speaker* (sum coded: SOV-Speaker = 1, OSV-Speaker = −1) and *Session* (sum coded: Training session = 1, EEG session = −1). A full random effects structure was implemented that included random intercepts for every participant and item as well as random slopes by participant for *Structure*, Speaker, *Session* and the interaction of all factors (Barr, Levy, Scheepers, & Tily, 2013; but see Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017). Possible main effects and interactions were evaluated using Type II Wald chi-square tests. Full model summaries are presented in the supplementary material.

### 2.6. *Open science statement*

Study materials and analysis are publicly available in a repository (https://osf.io/xez4d/). The conditions of our ethics approval and consent procedures do not permit public archiving of anonymized study data. However, this data will be released unconditionally on request to the corresponding author. No part of the study procedure or analysis was pre-registered prior to research being conducted. We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

## 3. Results

### 3.1. *Behavioral results*

Performance in the comprehension task is displayed in Fig. 1. With regards to error rate, the logit mixed-effect model revealed a main effect of *Structure*, $\chi^2(1) = 64.680$, $p < .001$, with increased error-rates for OSV structures compared to SOV structures ($\hat{\beta} = .42$, 95%-CI = [.31 .52]), and a main effect of *Session*, $\chi^2(1) = 44.567$, $p < .001$, with overall decreased error rates in the EEG session compared to the training session ($\hat{\beta} = -.38$, 95%-CI = [−.49 −.27]). There was no main effect of *Speaker* and no significant interactions involving either *Structure* or *Speaker*. The linear mixed-effects model for the analysis of the reaction times revealed only a main effect of *Session*, $\chi^2(1) = 46.121$, $p < .001$, with faster reaction times in the EEG session compared to the training session ($\hat{\beta} = 80.05$, 95%-CI = [54.98 105.12]). There was no main effect of *Structure* and

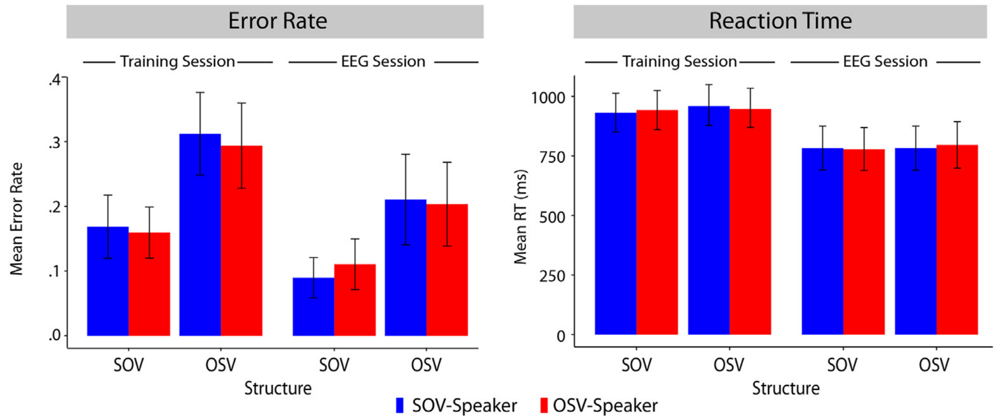


**Fig. 1 – Performance in the comprehension task. Error rates and reaction times are shown for different syntactic structures (SOV, OSV) and speaker condition (SOV-Speaker, OSV-Speaker) both in the training session and the EEG session. Error bars depict the standard error of the mean.**

*Speaker* and no interactions involving any of these factors. These results show that performance improved from the Training to the EEG session and that the comprehension question was more difficult for OSV structures than for SOV structures.

### 3.2. *ERP results*

#### 3.2.1. *Early time window: 300–500 msec*

A non-parametric cluster-based permutation test on the ERPs of the determiner of the first noun-phrase of the SOV and OSV sentence between 300 and 500 msec revealed a significant main effect of *Structure* ($p < .001$, see Fig. 2). This effect corresponded to a positive cluster between 312 and 500 msec. The cluster showed an increased positivity for OSV sentences compared to SOV sentences (see also Fig. 2B left). There were no significant negative clusters. Permutation tests were also conducted to test for the main effect of *Speaker* and the interaction between *Speaker* and *Structure*. These tests did not reveal any significant clusters in the early time window ($p > .05$). In summary, OSV structures elicited a greater positivity at centro-posterior electrodes than SOV structures.

#### 3.2.2. *Late time window: 600–900 msec*

Cluster-based permutation tests on the ERPs of the determiner of the first noun-phrase in the late time window revealed a main effect for *Structure*. There was a significant difference between OSV and SOV sentences ($p = .003$) with an increased positivity for OSV sentences compared to SOV sentences (Fig. 2B right). This effect was observed between 600 and 896 msec and was most pronounced on parietal electrodes. A similar analysis on the main effect of *Speaker* did not reveal any significant clusters ($p > .05$). The cluster-based permutation test on the interaction of *Speaker* and *Structure*, that is the difference between OSV and SOV structures for the OSV-Speaker condition compared to the difference between OSV structure and SOV structure for the SOV-Speaker condition, revealed a significant effect ($p = .038$). This effect was observed from 664 to 720 msec and showed a broad, central distribution (see Fig. 3 for single condition ERPs and Fig. 4 for difference waves depicting the difference between OSV and SOV structures for each speaker).

In order to follow-up on the interaction effect, we calculated step-down cluster-based permutation tests in the time window of the interaction effect in order to compare the effects of speaker identity on the processing of SOV and OSV structure sentences respectively. For OSV structure sentences there was a significant difference between speaker conditions, with an increased positivity when OSV structure sentences were produced by the SOV-Speaker compared to when OSV structure sentences were produced by the OSV-Speaker ($p = .04$). This effect was prominent over centro-posterior electrodes. For SOV structure sentences, only a marginal significant difference between speakers was observed, with a larger positivity for the OSV-Speaker compared to the SOV-Speaker ($p = .072$). This trending effect was prominent over posterior electrodes.

A further step-down analysis on the interaction effect was conducted in order to compare the effects of syntactic structure for each speaker separately (see Fig. 4). For the SOV-Speaker condition there was a significant difference between OSV and SOV sentences ($p = .005$), with an increased positivity for OSV sentences compared to SOV sentences that was prominent at posterior electrodes (Fig. 4B left). For the OSV-Speaker condition, however, there was a significant difference between OSV and SOV sentences ($p = .048$) with an increased negativity for OSV structures compared to SOV structures that was prominent over frontal electrodes (Fig. 4B right). There was only a marginal significant difference between OSV and SOV sentences related to an increased positivity for OSV structures compared to SOV structures in the OSV-Speaker condition at posterior electrodes ($p = .079$).

In summary, OSV sentences elicited an increased positivity compared to SOV sentences at posterior electrode sites. Importantly, a significant interaction effect was observed in the late time window. Responses to SOV and OSV structures were found to be modulated as a function of speaker identity.

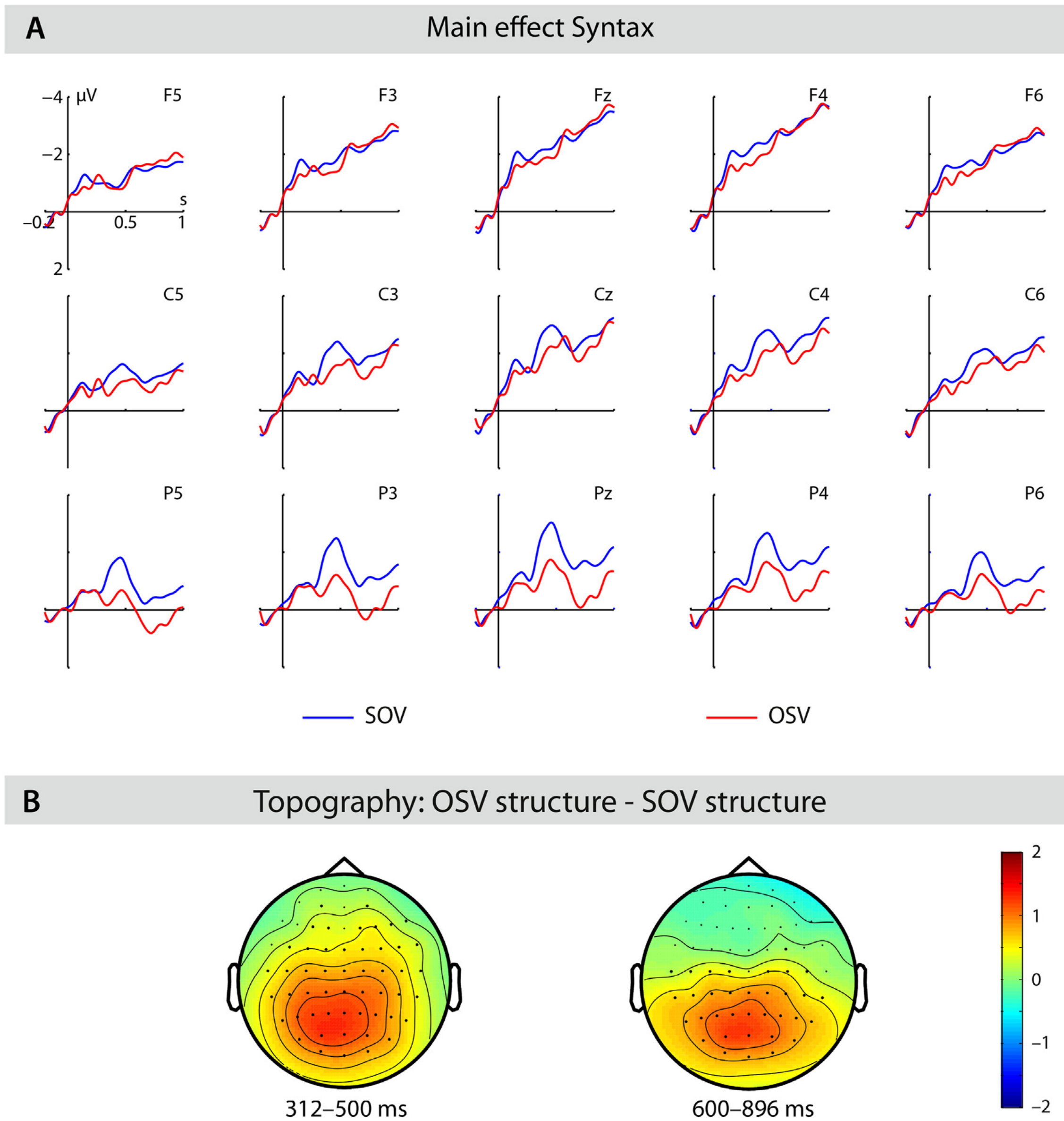


**Fig. 2 – Main effect of syntactic structure. (A) Event-related potentials related to the determiner of the first noun phrase of SOV and OSV sentences. (B) Topographical distribution of the difference between OSV and SOV sentences in the early and late cluster (312–500 msec, 600–896 msec).**

OSV sentences elicited an increased positivity when produced by the SOV-Speaker compared to the OSV speaker. A different pattern was observed for SOV sentences, which elicited an increased posterior positivity when produced by OSV-Speaker compared to the SOV-Speaker (albeit only marginal significant). Furthermore, the effect of syntactic structure modulates as a function of speaker identity. For the SOV-Speaker condition, a posterior positivity was observed when comparing difficult OSV sentences to easy SOV sentences. This effect resembled the canonical P600 effect. In contrast to this finding, a frontal effect was present for the OSV-Speaker condition, which showed an increased negativity for OSV structures compared to SOV structures. Furthermore, there was no posterior positivity for this comparison in case of the OSV-Speaker condition.

## 4. Discussion

The present study found an effect of speaker-specific expectations on neurophysiological correlates of syntactic processing during sentence comprehension. Interestingly, this interaction was only observed in a late time window but not in an early window where only general effects of syntactic processing were observed. Furthermore, we found differential effects depending on speaker identity. For the SOV-Speaker condition where a SOV structure was expected for a given sentence, we observed a posterior positivity that was increased for difficult OSV structures compared to easy SOV structures. The timing and topographical distribution of this effect resembled the P600 effect. Crucially, for the OSV-Speaker condition where an OSV structure was expected for

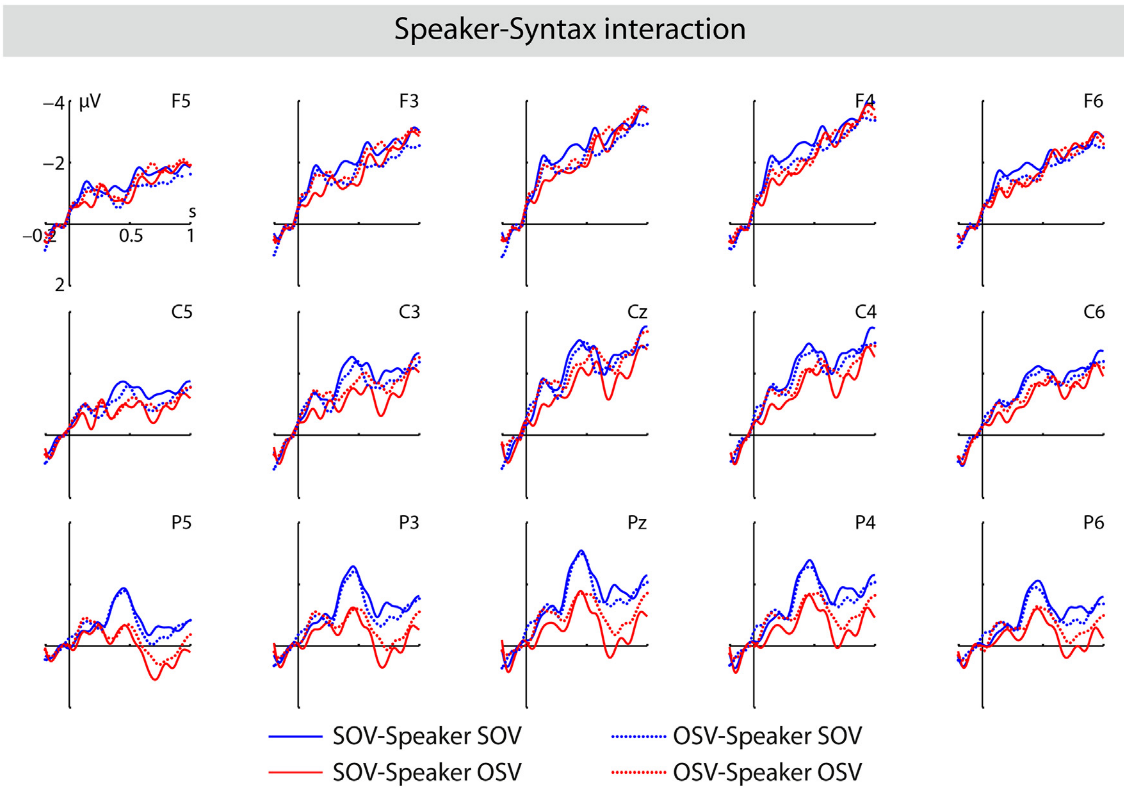


**Fig. 3 – Interaction effect of syntactic structure and speaker identity. Event-related potentials of the determiner of the first noun phrase for all four experimental conditions involved in the interaction effect. SOV structures are shown in blue and OSV structures are shown in red. The SOV-Speaker conditions are depicted as solid lines and the OSV-Speaker conditions are depicted as dashed lines.**

a given sentence, no such P600 but a frontal effect was observed. Instead, an increased negativity for OSV compared to SOV structures (or an increased positivity for SOV compared to OSV structures) was found. In summary, the present findings indicate that speaker-specific information is integrated with syntactic information only at a late stage of processing. Additionally, in line with previous neuroimaging findings (Kroczek & Gunter, 2020), the present results suggest that listeners process language input both with respect to a particular speaker-specific language use and a more global, population-wide distribution of language use.

Object-initial sentences in the SOV-Speaker condition were both difficult and unexpected. This is in line with the distribution of syntactic structures in everyday German language use, where object-initial structures are far more infrequent than subject-initial structures (Bader & Häussler, 2010). When presented without a licensing context, such object-initial structures elicit a P600 effect in comparison to easier subject-initial structures (Haupt, Schlesewsky, Roehm, ). The SOV-Speaker condition in the present study showed a similar P600-effect. The P600 has been related to syntactic processing in general (Hagoort, Brown, & Groothusen, 1993) and more specifically to syntactic reanalysis or repair (Friederici, 2002; Friederici, Hahne, & Mecklinger, 1996) as well as syntactic integration (Kaan, Harris, Gibson, & Holcomb, 2000). Furthermore, it has been hypothesized that the P600 reflects integration mechanisms which might reflect the effort related to constructing, reorganizing, and updating a mental representation of what is being communicated (Brouwer & Hoeks, 2013). It should be further noted, that the P600 has been linked to more domain general functions related to attentional processing and stimulus saliency in linguistically deviant stimuli (often referred to as the P600-as-P3 hypothesis, Sassenhagen & Bornkessel-Schlesewsky, 2015; Sassenhagen et al., 2014). In line with the P600-as-P3 hypothesis, the present results show that OSV sentences lead to an increased posterior positivity when they were presented in the SOV-Speaker condition, i.e., the condition where OSV sentences were infrequent and unexpected. It is therefore plausible to assume that in case of the SOV-Speaker condition, processing of OSV structures was more effortful and related to increased processing costs in comparison to the OSV-Speaker condition.

Interestingly, there was no pronounced P600 effect in the OSV-Speaker condition. Here, the difficult OSV structure was expected due to speaker identity, whereas the easy SOV structure was unexpected. It should be noted, that the diminished P600 effect in the OSV-Speaker condition could be driven both by the reduced positivity to OSV structures and an increased positivity to SOV structures. This indicates that processing of both structures in the late stage was influenced by expectancy (Sassenhagen et al., 2014). Similarly, previous studies found a modulation of the P600 components for various kinds of linguistic and non-linguistic information,

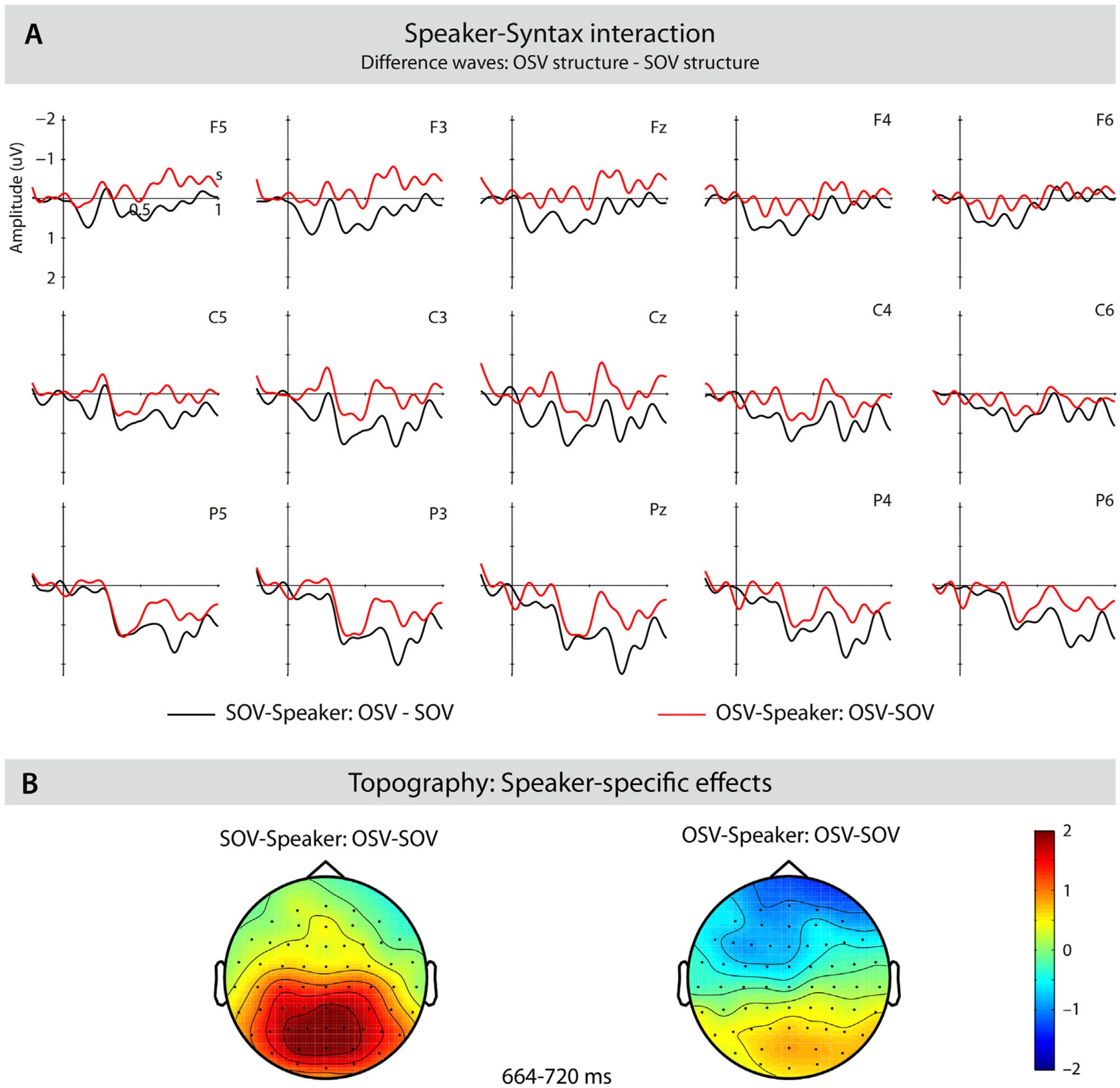


**Fig. 4 — *Effect of syntactic structure as a function of speaker identity.* (A) Difference waves showing the difference between ERPs related to OSV and SOV structure for the SOV-Speaker condition *in black and the OSV-Speaker* condition in red. (B) Topographical distribution of the difference between OSV and SOV structure for the SOV-Speaker (left) and OSV-Speaker (right) in a time window of 664–720 msec.**

such as gestures (Holle et al., 2012), pragmatic information like irony (Regel et al., 2010), but also speaker-related information such as accent and language style (Hanulíková et al., 2012; Viebahn, Ernestus, & McQueen, 2017). Furthermore, previous studies found no P600 effects when object-initial structures were licensed by contextual information (i.e., by dative object-experiencer verbs, Haupt et al., 2008). In analogy with these findings, speaker identity can be considered as contextual information that can be used to license particular structures. The present findings suggest that the P600 is sensitive to speaker-specific expectations based on language use (speaker context), thereby indicating that processing of an expected structure is less effortful than an unexpected structure (see also Fine et al., 2013). This suggests that the P600 indexes the integration of linguistic but also extra-linguistic information.

Crucially, we also observed a frontal ERP effect when comparing SOV and OSV structures in the OSV-Speaker condition. Recently, frontal ERP effects in sentence comprehension have been related to the shift from one internal mental model to another, when model predictions are violated (Brothers, Wlotko, Warnke, & Kuperberg, 2020; Kuperberg, Brothers, & Wlotko, 2020). Please note, that the shift described by the Kuperberg group has been related to a frontal positivity, while the present results showed a frontal negativity. In the present study, the ERPs of the difficult OSV structures were more negative compared to the easy SOV structures. In terms of expectancy, however, it might be more helpful to compare the unexpected SOV structure to the expected OSV structure (in case of the OSV-Speaker condition). With this comparison in mind, the present results show an increased positivity when an unexpected structure is processed that does not fit a listener's expectations based on speaker-specific language use. One could therefore speculate that an unexpected structure might trigger the shift from a local speaker-specific model to a global model of syntactic language use (e.g., general language priors). Importantly, SOV structures should be unexpected in the local OSV-Speaker model, but expected in a global model of German language use. Note, that such a model shift should only be observed in case of the OSV-Speaker condition where the local model deviates from the global model but not for the SOV-Speaker condition where the local and global model are similar. Conflicting model predictions for the SOV-Speaker condition might therefore trigger integration or re-analyses processes rather than model shifts. This notion is further supported by recent evidence that could show increased adaptation effects related to unexpected syntactic structures (Fine et al., 2013; Jaeger & Snider, 2013; Kroczek & Gunter, 2017).

The present findings also indicate that speaker-specific expectations are integrated only at a late processing stage, while early processing is solely driven by syntactic processing. This is supported by a fMRI study, where syntactic structure and speaker-specific expectancy processing lead to distinct neural activation, namely a left-lateralized fronto-temporal network for syntactic processing and a right-dominant fronto-parietal network that was related to speaker-specific syntactic expectations (Kroczek & Gunter, 2020). It needs to be acknowledged that the current results show an early, posterior positivity related to the syntactic manipulation, while other ERP studies typically report (left) anterior negativities for word orders that deviate from the more frequently used structure (Rösler, Pechmann, ; Schlesewsky, Bornkessel, & Frisch, 2003). Furthermore, other studies have reported a biphasic N400–P600 pattern when sentences were disambiguated towards difficult object-initial compared to easy subject-initial sentences (Haupt et al., 2008; Holle et al., 2012). Please note, that in the present experiment we did not present ambiguous sentence structures, which were disambiguated at the sentence-final position. In contrast, we used non-ambiguous sentences where the syntactic structure became clear at the first noun phrase of the sentence. While these differences should be investigated in future studies, they do not challenge our finding of an early syntactic effect that was unaffected by speaker expectations, and a late interaction of syntactic information and speaker-specific expectations.

It should further be noted that the 24-h delay between the training and the EEG session might have introduced memory effects, which could be related to either attenuation or consolidation of speaker-specific expectations (Feld & Born, 2017). The present study design does not allow testing potential memory effects. Future studies should specifically target the role of memory consolidation in the formation of speaker-specific expectations. In this respect, it is worth mentioning that a previous study found speaker-specific expectations being rapidly re-activated within a short period of exposure to the speakers, even after a period of nine months (Kroczek & Gunter, 2017). This finding seems to speak against attenuation effects.

In summary, the present EEG study investigated neurophysiological correlates of speaker-specific expectations in syntactic processing. Speaker-specific effects were only observed in a late processing stage, while early processing was solely driven by syntactic processing. Importantly, two distinct ERP patterns were observed for speaker-specific processing of sentence structure. A P600 effect was related to the effort of integrating conflicting information, when a speaker's language use was similar to the typical everyday language use. Furthermore, a frontal effect was observed when conflicting information between local and global representations of language use induced a model shift in the listeners. The study supports the notion of speaker-specific language expectations that are processed in addition to general language priors.

## Author contributions

L.K. and T.G. designed research, L.K. created stimuli, supervised data acquisition and analyzed data. L.K. and T.G. wrote the paper.

## Open practices

The study in this article earned an Open Materials badge for transparent practices. Materials for this study can be found at https://osf.io/xez4d/.

## Competing financial interests

The authors declare no conflict of interest.

## Acknowledgments

We are grateful to Angela D. Friederici for her kind support of the research described here and her helpful discussions during the preparation of this paper. We thank Kerstin Flake for her help in assembling the figures.

## Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2021.04.017.

## REFERENCES

Bader, M., & Häussler, J. (2010). Word order in German: A corpus study. *Lingua, 120*(3), 717–762. https://doi.org/10.1016/j.lingua.2009.05.007

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review, 113*(4), 787–821. https://doi.org/10.1037/0033-295X.113.4.787

Brothers, T., Dave, S., Hoversten, L. J., Traxler, M. J., & Swaab, T. Y. (2019). Flexible predictions during listening comprehension: Speaker reliability affects anticipatory processes.

*Neuropsychologia, 135*, 107225. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2019.107225

Brothers, T., Wlotko, E. W., Warnke, L., & Kuperberg, G. R. (2020). Going the extra mile: Effects of Discourse context on two late Positivities during language comprehension. *Neurobiology of Language, 1*(1), 135–160. https://doi.org/10.1162/nol_a_00006

Brouwer, H., & Hoeks, J. C. J. (2013). A time and place for language comprehension: Mapping the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience, 7*(November), 1–12. https://doi.org/10.3389/fnhum.2013.00758

Brown-Schmidt, S., Yoon, S. O., & Ryskin, R. A. (2015). People as contexts in conversation. In *Psychology of learning and motivation – Advances in research and theory* (Vol. 62). Elsevier Ltd. https://doi.org/10.1016/bs.plm.2014.09.003

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491–505. https://doi.org/10.1111/j.1469-8986.2007.00531.x

Feld, G. B., & Born, J. (2017, June 1). Sculpting memory during sleep: Concurrent consolidation and forgetting. *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2017.02.012. Elsevier Ltd.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS One, 8*(10), Article e77661. https://doi.org/10.1371/journal.pone.0077661

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression (second)*. Thousand Oaks, CA: SAGE Publications.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*(2), 78–84. https://doi.org/10.1016/S1364-6613(00)01839-8

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews, 91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Friederici, A. D. (2018). *Language in our brain* (Vol. 1). Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/9780262036924.001.0001

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning Memory and Cognition, 22*(5), 1219–1248. https://doi.org/10.1037/0278-7393.22.5.1219

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience, 12*(4), 556–568.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (sps) as an erp measure of syntactic processing. *Language and Cognitive Processes, 8*(4), 439–483. https://doi.org/10.1080/01690969308407585

Hale, J. (2001). A probabilistic earley parser as a psycholinguistic model. In *Second meeting of the North American Chapter of the association for computational linguistics on language technologies* (pp. 1–8). https://doi.org/10.3115/1073336.1073357. Pittsburgh, Pennsylvania.

Hanulíková, A., van Alphen, P. M., van Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience, 24*(4), 878–887. https://doi.org/10.1162/jocn_a_00103

Haupt, F. S., Schlesewsky, M., Roehm, D., Friederici, A. D., & Bornkessel-Schlesewsky, I. (2008). The status of subject-object reanalyses in the language comprehension architecture. *Journal of Memory and Language, 59*(1), 54–96. https://doi.org/10.1016/j.jml.2008.02.003

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics, 48*(4), 865–892. https://doi.org/10.1515/LING.2010.027

Hellbernd, N., & Sammler, D. (2018). Neural bases of social communicative intentions in speech. *Social Cognitive and Affective Neuroscience, 13*(6), 604–615. https://doi.org/10.1093/scan/nsy034

Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience, 19*(7), 1175–1192.

Holle, H., Obermeier, C., Schmidt-Kassow, M., Friederici, A. D., Ward, J., & Gunter, T. C. (2012). Gesture facilitates the syntactic analysis of speech. *Frontiers in Psychology, 3*, 74. https://doi.org/10.3389/fpsyg.2012.00074

Huettig, F., Mani, N., & Huettig, F. (2015). Is prediction necessary to understand language? Probably not. *Language, Cognition, and Neuroscience, 3798*(January), 1–26. https://doi.org/10.1080/23273798.2015.1072223

Jaeger, T. F., & Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition, 127*(1), 57–83. https://doi.org/10.1016/j.cognition.2012.10.013

Jung, T., Makeig, S., Humphries, C., Lee, T., McKeown, M. J., Iragui, I., et al. (2000). Removing electroencephalographic aretfacts by blind source separation. *Psychophysiology, 37*(2), 163–178. https://doi.org/10.1111/1469-8986.3720163

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15*(2), 159–201. https://doi.org/10.1080/016909600386084

Kroczek, L. O. H., & Gunter, T. C. (2017). Communicative predictions can overrule linguistic priors. *Scientific Reports, 7*(1), 17581. https://doi.org/10.1038/s41598-017-17907-9

Kroczek, L. O. H., & Gunter, T. C. (2020). Distinct neural networks relate to common and speaker-specific language priors. *Cerebral Cortex Communications*. https://doi.org/10.1093/texcom/tgaa021

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience, 32*(1), 12–35. https://doi.org/10.1162/jocn_a_01465

Kuperberg, G. R., & Jaeger, T. F. (2015). What do we mean by prediction in language comprehension? *Language Cognition & Neuroscience, 3798*(January), 1–70. https://doi.org/10.1080/23273798.2015.1102299

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in linear mixed effects models*.

Lattner, S., & Friederici, A. D. (2003). Talker's voice and gender stereotype in human auditory sentence processing – Evidence from event-related brain potentials. *Neuroscience Letters, 339*(3), 191–194. https://doi.org/10.1016/S0304-3940(03)00027-2

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011. https://doi.org/10.1155/2011/156869

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34*(6), 739–773. https://doi.org/10.1006/jmla.1995.1033

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from https://www.r-project.org/.

Regel, S., Coulson, S., & Gunter, T. C. (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain Research, 1311*, 121–135. https://doi.org/10.1016/j.brainres.2009.10.077

Rösler, F., Pechmann, T., Streb, J., Ro, B., & Hennighausen, E. (1998). Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory and Language, 176*(38), 150–176.

Sassenhagen, J., & Bornkessel-Schlesewsky, I. (2015). The P600 as a correlate of ventral attention network reorientation. *Cortex, 66*, A3–A20. https://doi.org/10.1016/j.cortex.2014.12.019

Sassenhagen, J., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2014). The P600-as-P3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain and Language, 137*, 29–39. https://doi.org/10.1016/j.bandl.2014.07.010

Schlesewsky, M., Bornkessel, I., & Frisch, S. (2003). The neurophysiological basis of word order variations in German. *Brain and Language, 86*(1), 116–128. https://doi.org/10.1016/S0093-934X(02)00540-0

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia, 56*(1), 289–301. https://doi.org/10.1016/j.neuropsychologia.2014.02.002

Van Berkum, J. J. A. (2008). Understanding sentences in context. *Current Directions in Psychological Science, 17*(6), 376–380. https://doi.org/10.1111/j.1467-8721.2008.00609.x

Viebahn, M. C., Ernestus, M., & McQueen, J. M. (2017). Speaking style influences the brain's electrophysiological response to grammatical errors in speech comprehension. *Journal of Cognitive Neuroscience, 29*(7), 1132–1146.

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data – A practical approach. *Journal of Neuroscience Methods, 250*, 34–46. https://doi.org/10.1016/j.jneumeth.2014.08.002.Digital